M2 CRYPTO
TP THÉORIE ALGORITHMIQUE DES NOMBRES

CHRISTOPHE RITZENTHALER

1. Réseaux

Exercice 1. Implémenter l'algorithme de Gauss. On devra utiliser la fonction `while`

Algorithm 1: Algorithme de Gauss

input : Une base ordonnée (u, v) avec $q(u) \leq q(v)$
output: Une base réduite du réseau
repeat
  $x = \lfloor (u, v)/q(u)^2 \rceil$
  $r = v - xu$
  $v = u$
  $u = r$
until $q(u) \geq q(v)$;
return (v, u)

d'où une légère modification du programme. Pour les opérations, on pourra se servir de `vector()`, pour lesquels on a une fonction `norm` et un produit scalaire avec $*$.

Exercice 2. Implémenter l'algorithme LLL du cours : on stockera les vecteurs de la base en ligne dans une matrice A. On aura besoin d'un algorithme pour calculer les coefficients μ_{ij} de GSO : on utilisera `A.gram_schmidt()` si on veut gagner du temps. L'algorithme pour la proprification est donné ci-dessous. On pourra utiliser les fonctions

Algorithm 2: proprification de b_t par rapport à $(b_1, \ldots, b_{t-1})$ (déjà propres)

input : Une base $(b_1, \ldots, b_n)$ avec $b_1, \ldots, b_{t-1}$ propres et b_t
output: une base avec $b_1, \ldots, b_t$ propres
Calculer la GSO des b_i, stocker dans une matrice (μ_{ij})
$i = t$
for $j = i - 1$ *to* 1 **do**
  $x_j = \lfloor \mu_{ij} \rceil$
  $b_i = b_i - x_j b_j$
  $\mu_{ij} = \mu_{ij} - x_j$
  for $k = 1$ *to* $j - 1$ **do**
    $\mu_{ik} = \mu_{ik} - x_j \mu_{jk}$
return (b_i).

`ncols(),set_row()` et tester le fonctionnement (par exemple sur la matrice $\begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 4 & 3 & 3 \end{bmatrix}$) en comparant le résultat à la sortie de `M.LLL()` (attention il n'y a pas unicité du résultat en général) et en vérifiant que le résultat de sortie est une base réduite.

Exercice 3. On utilise ici LLL pour trouver des relations entre réels [1]. Soient $x_1, \ldots, x_d$ des réels. On cherche $n_1, \ldots, n_d$ des entiers tels que $|\sum_{i=1}^{d} n_i x_i|$ est petite. Donnons l'idée de la méthode. Soit L le réseau engendré par les lignes b_i de la matrice suivante

$$B = \begin{bmatrix} \lfloor Cx_1 \rceil & 1 & 0 & \ldots & 0 \\ \lfloor Cx_2 \rceil & 0 & 1 & \ldots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ \lfloor Cx_d \rceil & 0 & \ldots & 0 & 1 \end{bmatrix}$$

Le premier vecteur de sortie de LLL est un vecteur court de la forme

$$n_1 b_1 + n_2 b_2 + \ldots + n_d b_d = (\sum_{i=1}^{d} n_i \lfloor Cx_1 \rceil, n_1, \ldots, n_d).$$

Si C est grand alors cela implique que $|\sum n_i x_i|$ est petite, ce qu'on voulait et les autres coordonnées donnent les n_i permettant d'obtenir cela. On va tester cette stratégie sur deux exemples.

(1) On cherche le polynôme minimal de $\sqrt{2} + \sqrt{3} + \sqrt[3]{3}$. On travaillera avec des réels `RealField(400)` et $C = 10^{100}$.

(2) On cherche une relation linéaire entre π et des expressions de la forme $\sum_{k=0}^{\infty} \frac{1}{16^k(8k+i)}$ avec $i = 1, \ldots, 8$ (Formule de Bailey, Borwein et Plouffe). On prendra $C = 10^{600}$, $k = 1000$ dans `RealField(2000)`.

2. Primalité - Factorisation

Exercice 4. Essayer la méthode $p-1$ de Pollard sur $F_6 = 2^{35} + 1$.

Exercice 5. Essayer la méthode ECPP pour prouver la primalité de $n = 2000003$. On pourra par exemple construire une courbe elliptique de la forme $E : y^2 = x^3 + B$ pour laquelle on sait facilement évaluer le nombre de points.

Exercice 6. On va étudier la factorisation par crible linéaire, quadratique (QS) et crible du corps de nombres (NFS) dans un cas particulier.

Racines carrées modulo un entier.
Soit $n = p_1 \ldots p_s$ un entier composé, les p_i étant premiers impairs distincts.

1. Quelles sont les racines carrées de 1 modulo p_i ?
2. Montrer que l'équation $x^2 = 1 \bmod n$ admet 2^s solutions dans $\mathbb{Z}/n\mathbb{Z}$.
3. Soit x une telle solution différente de ± 1. Montrer que $\gcd(x-1, n)$ et $\gcd(x+1, n)$ sont des facteurs non triviaux de n.
4. Plus généralement, soient x, y, n des entiers tels que $x^2 = y^2 \bmod n$, et $x \neq \pm y \bmod n$. Montrer que $\gcd(x-y, n)$ et/ou $\gcd(x+y, n)$ fournit un facteur non trivial de n.

Les méthodes de crible (RS, QS, NFS) visent toutes à trouver une relation de ce type. Elles diffèrent les unes des autres par la manière de trouver une telle relation. Nous illustrons sur des exemples numériques.

Dans toutes les méthodes, on suppose que l'on a déjà "retiré" de n les éventuels petits

1. Son incarnation ultime est l'inverseur de Plouffe qui n'existe plus mais qui a un analogue `http://isc.carma.newcastle.edu.au/`

facteurs premiers, c.-à-d. les facteurs premiers inférieurs à une borne que l'on s'est donnée à l'avance. Ce peut être fait de manière naïve si la borne est effectivement "petite" (polynomiale en $\log n$).

Crible linéaire.
Le but de l'exercice est de trouver un facteur non trivial de $n = 7081$. On pose $S = \{-1, 2, 3\}$ que l'on appelle la *base de facteurs*. On vérifiera que n n'est pas divisible par les premiers de la base de facteurs.

1. Dans cette question on cherche des entiers x tels que $\pm x^2 \pmod n$ soient S-lisses[2]. Montrer que $x = 4486, 1857, 2645$ conviennent. On a par exemple

$$4486^2 \equiv -2 \cdot 3 \pmod n$$

2. Montrer qu'en combinant (multiplicativement) des lignes, on obtient une relation de la forme $x^2 = y^2 \bmod n$.

Si l'on prend une combinaison quelconque des lignes du tableau, il y a peu de chances de tomber sur une relation du type $X^2 = Y^2 \bmod n$. La suite de l'exercice montre que ce problème combinatoire peut être reformulé en un problème d'algèbre linéaire sur $\mathbb{F}_2$, et (donc) résolu efficacement.

Pour chaque x et chaque signe ϵ, tel que $y = \epsilon x^2 \pmod n$ est friable, on forme le vecteur v dans $\mathbb{F}_2^3$ avec en première position 0 si le signe est positif et 1 sinon, suivi de la valuation modulo 2 de y en 2 et en 3. Par exemple

$$v_1 \quad = \quad (1, 1, 1)$$

3. Montrer qu'il existe une combinaison linéaire nulle non triviale des v_i.
4. En déduire une factorisation de n.

Crible quadratique.
Le but l'exercice est de trouver un facteur non trivial de $n = 24961$. Comme indiqué précédemment, on cherche toujours une relation de la forme $X^2 = Y^2 \bmod n$, mais la manière de l'obtenir est différente.

On pose $m = \lfloor \sqrt{24961} \rfloor = 157$, et

$$q(x) = (x + m)^2 - n = (x + 157)^2 - 24961.$$

L'idée est de trouver des x tels que $q(x)$ (positif ou négatif) soit lisse. L'heuristique est que $q(x)$ est petit par rapport à n si x l'est, et a plus de chances d'être lisse qu'un nombre quelconque. Ici, on prend 23 comme borne de lissité. La base de facteurs est donc $\{-1, 2, 3, 5, 7, 11, 13, 17, 19, 23\}$. La présence du "$-1$" permet d'accepter les $q(x)$ lisses négatifs. Autrement dit, on va chercher des x tels que $q(x)$ (de signe quelconque) n'ait que des facteurs premiers ≤ 23.

1. Soit x convenable, et p un facteur premier de $q(x)$. Montrer que n est un carré modulo p.
2. En déduire que, pour tout x convenable, $q(x)$ ne contient jamais les facteurs premiers $7, 11, 17, 19$.

2. C'est-à-dire dont les facteurs premiers sont dans S.

On prend finalement $S = \{-1, 2, 3, 5, 13, 23\}$ comme *base de facteurs*. Réduire ainsi la base de facteurs peut sembler artificiel, mais l'intérêt va apparaître après la prochaine question.

3. Pour $x = 0, 1, -1, 2, -2, \ldots, 6, -6$, tester si $q(x)$ est S-lisse. Si oui, poser :
— $a_x = x + m$,
— $b_x = q(x) = (-1)^{e_{x,0}} 2^{e_{x,1}} 3^{e_{x,2}} 5^{e_{x,3}} 13^{e_{x,4}} 23^{e_{x,5}}$, (NB : $e_{x,0} \in \{0, 1\}$)
— $e_x = (e_{x,0}, e_{x,1}, e_{x,2}, e_{x,3}, e_{x,4}, e_{x,5}) \in \{0, 1\} \times \mathbb{N}^5$
— $v_x = (v_{x,0}, \ldots v_{x,5})$ le vecteur binaire où $v_{x,i} = e_{x,i} \bmod 2$
et conserver le tout dans un tableau.

4. Montrer qu'il existe au moins une combinaison linéaire binaire nulle des vecteurs v_x collectionnés.

5. Quel est l'intérêt d'avoir réduit la base de facteurs après la question 2 ?

6. Vérifier, à titre d'exemple, que $v_{-1} + v_4 + v_{-6} = 0$.

7. En déduire que $b_{-1} b_4 b_{-6}$ est un carré S-lisse, et qu'il est égal à $a_{-1}^2 a_4^2 a_{-6}^2$ modulo n (on observera que par construction $a_x^2 = b_x \bmod n$)

8. En déduire une écriture $X^2 = Y^2 \bmod n$, et deux facteurs non triviaux de 24961.

Crible algébrique.
Le crible algébrique "général" est la méthode la plus efficace connue pour factoriser des nombres quelconques. Le record de factorisation est un nombre de 768 bits, réalisé en 2010 (après plus d'un an de calcul).

Cela dit, on peut faire mieux dans le cas de nombres "spéciaux", de la forme $x^s \pm e$ où e est petit. La variante du crible s'appelle crible algébrique "spécial". L'exercice, inspiré de l'article de J. M. Pollard, "Factoring with cubic integers", montre la factorisation de nombre de la forme $x^3 + 2$, en l'illustrant avec $x = 2^{43}$. Cela permet au passage de factoriser le septième nombre de Fermat $F_7 = 2^{2^7} + 1$, puisque $2F_7 = (2^{43})^3 + 2$. En particulier, cela prouve que F_7 n'est pas premier.

1. Montrer que le polynôme $X^3 + 2$ est irréductible dans $\mathbb{Q}[X]$ (on pourra utiliser le critère d'Eisenstein).

On considère l'extension algébrique $K = \mathbb{Q}(\alpha)$ avec $\alpha^3 + 2 = 0$. On a $K = \{a + b\alpha + c\alpha^2 : a, b, c \in \mathbb{Q}\}$, et c'est une extension de degré 3 de $\mathbb{Q}$. La norme (produit des conjugués dans une clôture algébrique de K) d'un élément $a + b\alpha + c\alpha^2 \in K$ est égale à

$$\mathrm{Norm}(a + b\alpha + c\alpha^2) = a^3 - 2b^3 + 4c^3 + 6abc.$$

C'est une fonction multiplicative sur K. On construit cette extension avec la fonction `K.<a> = NumberField(x^3 + 2)`.

On note $\mathcal{O}$ l'*anneau des entiers* de K, c.-à-d. l'ensemble des éléments de K racines d'un polynôme unitaire à coefficients dans $\mathbb{Z}$. On note U l'ensemble des unités, c.-à-d. l'ensemble des inversibles de $\mathcal{O}$.

Un nombre $x \in \mathcal{O}$ est dit "premier" s'il n'est pas dans U et si pour toute factorisation $x = yz$ dans $\mathcal{O}$, on a y ou $z \in U$.

Nous admettrons les résultats suivants, qui peuvent être faux dans le cas général, mais

vrais pour l'exemple. C'est ce qui rend cet exemple agréable et pédagogique pour "débuter" dans le crible algébrique.

Théorème (1) $\mathcal{O} = \{a + b\alpha + c\alpha^2 : a, b, c \in \mathbb{Z}\}$ *(NB : en général on a seulement* $\mathcal{O} \supseteq \{\ldots\}$*). On peut vérifier cela avec* `K.integral_basis()`.
(2) $\mathcal{O}$ *est un anneau factoriel (faux en général). On peut vérifier cela avec la fonction* `K.is_unique_factorization_domain()`.
(3) U *est l'ensemble des éléments de* $\mathcal{O}$ *de norme égale à* ± 1 *(toujours vrai).*
(4) U *est un groupe multiplicatif engendré par* ± 1 *et* $\alpha - 1$*. On peut vérifier cela avec* `U=K.unit_group()` *puis* `U.gens_values()`.

Soit $x \in \mathcal{O}$. Le point (1) du théorème implique immédiatement $\mathrm{Norm}(x) \in \mathbb{Z}$.

1. Soit $x \in \mathcal{O}$ et $y \in \mathcal{O}$ un diviseur de x. Montrer que $\mathrm{Norm}(y)$ divise $\mathrm{Norm}(x)$.
2. En déduire que si $\mathrm{Norm}(x)$ est premier dans $\mathbb{Z}$, x est premier dans $\mathcal{O}$.

Soit p un nombre premier dans $\mathbb{Z}$, et soit x un diviseur premier de p dans $\mathcal{O}$.

3. Montrer que $\mathrm{Norm}(x) \in \{\pm p, \pm p^2, \pm p^3\}$.

Le résultat suivant, que nous admettrons, établit le type de factorisation des premiers de $\mathbb{Z}$ en facteurs premiers de $\mathcal{O}$. Sur un exemple, cela peut être vérifier avec `K.factor()`.

Théorème (1) *Si* $p = 2$ *ou* 3 *: p est le cube d'un premier de norme p (à unité près)*

$$2 = (\underbrace{-1}_{\in U})\ \underbrace{\alpha}_{Norme\ -2}{}^3 \quad ; \quad 3 = \underbrace{(1+\alpha)}_{\in U}\underbrace{(-1+\alpha)}_{Norme\ -3}{}^3$$

(2) *Si* $p = 6m + 1$ *et* -2 *est un cube modulo p : p est le produit de trois premiers de norme p (à unité près). Exemple :*

$$31 = (\underbrace{5 - 4\alpha + 3\alpha^2}_{\in U})(\underbrace{-1 - 2\alpha + \alpha^2}_{Norme\ 31})(\underbrace{-9 - 6\alpha + \alpha^2}_{Norme\ 31})(\underbrace{3 + \alpha^2}_{Norme\ 31})$$

(3) *Si* $p = 6m + 5$ *: il y a un facteur de norme p et un de norme* p^2*. Exemple :*

$$5 = (\underbrace{1 + \alpha^2}_{Norme\ 5})(\underbrace{1 + 2\alpha - \alpha^2}_{Norme\ 25})$$

(4) $p = 6m + 1$ *et* -2 *n'est pas un cube modulo p : p est un premier de norme* p^3*.*

On en vient maintenant au crible lui-même. On va chercher des entiers a, b tels que :
— $\mathrm{pgcd}(a, b) = 1$;
— $a + b2^{43}$ soit lisse dans $\mathbb{Z}$;
— et $a + b\alpha$ soit lisse dans $\mathcal{O}$, ce qui équivaut à ce que $\mathrm{Norm}(a + b\alpha) = a^3 - 2b^3$ soit lisse dans $\mathbb{Z}$.

On construit deux bases de facteurs :
— S_1 est l'ensemble des 500 premiers nombres premiers : $\{2, 3, 5, \ldots, 3571\}$; dans S_1, il y a :
  — 81 éléments de catégorie (2) du théorème ;
  — 252 éléments de catégorie (3) ;
  — 165 éléments de catégorie (4) ;

— S_2 est l'ensemble des éléments de $\mathcal{O}$ qui :
  — sont des facteurs premiers des éléments de S_1,
  — ne sont pas déjà dans S_1,
  — qui sont de norme $\pm p$ (voir question 5 pour la justification)
  auxquels on adjoint trois unités : $-1, 1+\alpha, (1+\alpha)^{-1} = -1+\alpha-\alpha^2$.

4. Soit p un nombre premier du cas (3) du théorème, et x un facteur premier de norme p^2. Montrer que x ne peut pas diviser $a+b\alpha$, $\text{pgcd}(a,b) = 1$.

L'objectif des questions 5 et 6 est de compter les éléments de S_2.

5. a. Montrer que $p = 2$ et 3 (cas (1) du théorème) fournissent chacun un nombre premier de $\mathcal{O}$ dans S_2.
   b. Montrer que tout p du cas (2) fournit trois nombres premiers de $\mathcal{O}$ dans S_2
   c. Montrer que tout p du cas (3) fournit un nombre premier de $\mathcal{O}$ dans S_2.
   d. Montrer que tout p du cas (4) ne fournit aucun nombre premier de $\mathcal{O}$ dans S_2.

6. En déduire que S_2 possède 497 éléments.

7. Montrer que $N(S_2) = \{\text{Norm}(x); x \in S_2\}$ possède 335 valeurs premières distinctes.

On note $\pi_1, \ldots, \pi_{497}$ les premiers de S_2.

La suite du crible fonctionne comme suit : on fait varier a et b dans les intervalles : $a \in [-4800, 4800]$, $b \in [1, 2000]$ pour trouver des nombres $a + b2^{43}$ S_1-lisses. On obtient des relations de la forme

$$a + b2^{43} = 2^{r_1}3^{r_2} \cdot \cdots \cdot 3571^{r_{500}}.$$

En pratique, on pourra tester cela avec $a \in [-48, 48]$ et $b \in [1, 20]$.

Pour chaque couple (a, b) obtenu, on regarde si la norme $\text{Norm}(a+b\alpha)$ est lisse sur $N(S_2)$. Si oui, on peut obtenir aussi une relation du type

$$a + b\alpha = (\pi_1^{s_1} \cdot \cdots \cdot \pi_{497}^{s_{497}})u, u \in U$$

En résumé, on a collecté des couples (a, b) tels que l'on ait simultanément :

$$\begin{aligned} a + b2^{43} &= \prod_{p \in S_1} p^{r_p} \quad \text{(collection A)} \\ a + b\alpha &= \prod_{\pi \in S_2} \pi^{s_\pi} \quad \text{(collection B)} \end{aligned}$$

L'idée suivante est d'obtenir des équations de lissité modulo $2F_7$:

$$a + b2^{43} = \prod_{p \in \ldots} p^{v_p} \pmod{2F_7}$$

8. Montrer que la collection A fournit naturellement une collection A' de telles équations.

Pour la collection B, il suffirait de "remplacer" formellement α par 2^{43}. Le point crucial

est que c'est possible grâce au lien étroit entre le corps de nombres introduit et le nombre à factoriser. En effet, rappelons que :

— le nombre à factoriser est $2F_7 = (2^{43})^3 + 2$;

— le polynôme irréductible de définition du corps de nombres est $X^3 + 2$.

9. Montrer que

$$\begin{array}{rccl} \phi & : & \mathcal{O} & \to & \mathbb{Z}/2F_7\mathbb{Z} \\ & & a + b\alpha + c\alpha^2 & \mapsto & a + b2^{43} + c2^{86} \end{array}$$

définit un morphismes d'anneaux.

10. En déduire que la collection B donne une collection B' d'équations de la forme

$$a + b2^{43} = \prod_{\pi \in S_2} \phi(\pi)^{s_\pi} \bmod 2F_7$$

Au bilan, chaque couple (a, b) adéquat fournit une équation pour la collection A' et une pour la collection B'. En les rapprochant 2 à 2, on voit que chaque couple (a, b) donne une équation de la forme :

$$\prod_{p \in S_1} p^{r_p} = \prod_{\pi \in S_2} \phi(\pi)^{s_\pi} \bmod 2F_7 \tag{1}$$

12. Montrer que si la collection des équations de la forme 1 contient suffisamment d'équations (combien ?), on peut exhiber une relation $X^2 = Y^2 \bmod 2F_7$.

L'avantage de cette méthode par rapport au crible quadratique est qu'elle permet de réduire la complexité. L'heuristique est que le polynôme de définition du corps est de degré 3, et ses coefficients sont de l'ordre de $(2F_7)^{1/3}$.

Pour que le crible spécial marche, il faut que le nombre à factoriser soit de la forme $r^e \pm s$ avec r, s petits. De plus, la racine modulo $n = 2F_7$ du polynôme de définition du corps (ici, cette racine est 2^{43}) doit être aussi racine d'un binôme de degré 1 à coefficients de l'ordre de $n^{1/3}$: c'est le cas ici puisque 2^{43} est racine du binôme $X - 2^{43}$ dont les coefficients sont de l'ordre de $(2F_7)^{1/3}$.

3. Méthode AGM

Exercice 7. Implémenter les fonctions thêta en genre 1 et vérifier sur des exemples les différentes relations du cours.

Exercice 8. On cherche à implémenter la méthode AGM pour le comptage de points sur une courbe elliptique ordinaire de la forme

$$E : y^2 + xy = x^3 + a_2x^2 + a_4x$$

sur $\mathbb{F}_q$ avec $q = 2^d$. On commencera par définir l'extension 2-adique non ramifiée de degré d, $\mathbb{Q}_q$, par `Q2.<a>=Qq(2^d,d+5)`, puis son corps résiduel F=`Q2.residue_field()`. Un générateur de $F = \mathbb{F}_q$ est `a0=F.0`.

Une des difficultés de l'implémentation est qu'il n'y a pas de fonction racine carrée implémentée pour $\mathbb{Q}_q$. Pour les $a \equiv 1 \pmod 8$, on pourra contourner la difficulté en prenant $\sqrt{a} = \exp(\log(a)/2)$. Il faudra aussi penser à compenser les pertes de précision qu'entraînent les racines carrées en relevant les valeurs avec `lift_to_precision()`.

On pourra finir en effectuant des comparaisons de temps entre la fonction de Sage et celle qu'on vient d'implémenter pour différentes valeurs de d.